Date: 2003-12-12

DRAFT version – 0.1

Attached Binary Object

NOTE: This document is a Working Draft that must be considered as "work in progress." It has not reached the stage of submission to the Environmental Data Standards Council.

This standard has been produced through the Environmental Data Standards Council (EDSC)

The Environmental Data Standards Council (EDSC) is a partnership among EPA, States and Tribal partners to develop and agree upon data standards for environmental information collection and exchange. The Council seeks to promote efficient sharing of environmental information between State, EPA and Tribal partners through the development of data standards. More information about the EDSC is available at www.epa.gov/edsc/.

Contents

Tables

Foreword

The Environmental Data Standards Council (EDSC) is a partnership among EPA, States and Tribal partners to develop and agree upon data standards for environmental information collection and exchange. The Council seeks to promote efficient sharing of environmental information between State, EPA and Tribal partners through the development of data standards. More information about the EDSC is available on the EDSC website at www.epa.gov/edsc/.

The Council identifies, prioritizes and pursues those areas where information exchange standards will provide the most value in achieving environmental results. The Council involves Tribes and Tribal Nations, and other state and federal agencies in the development of the standards and then provides the draft materials for general review. Business groups, non-governmental organizations, and other interested parties may then provide input and comment for Council consideration and standard finalization.

Introduction

Environmental information is a key tool in the effective management of our environmental resources and human health conditions. As a result, much effort goes into data acquisition, management, maintenance, exchange, and oversight. Greater access is the goal of many data consumers, and data managers. Providers invest significant resources meeting their requirements. In response, many data providers are improving access as they post usable copies of their environmental information on the web. These efforts are a vast improvement over previous conditions; however, there is a growing desire and need to both provide and receive data in a clearly defined and a uniform way. Data from multiple sources can then be aggregated and used without the inherent variations that exist between data sets across agencies.

Data exchange standards have been in place for years, e.g., health care, UPC codes for commodities, etc. The Attached Binary Object Data Standard is the next step to improve the exchange of environmental data across the nation. Implementers of the standard will map their own data to the nomenclature specified in the standard and develop the translation/export scripts. Once done, extractions and exchanges become routine, and no changes to their parent data management systems will be necessary.

Implementation of the standard will require some short-term investment as discussed: however, there will be short and longer-term benefits for users. These include:

- Data quality improvement—data that are being exchanged will be evaluated by the sender and the recipient—data issues will be discovered and corrected;
- Human intervention in the interchange of information machine-to-machine environments is reduced;
- Rework is reduced—cleaner data require less maintenance and correction;
- Public resources are conserved— if access is reliable and fast, data collected by another organization can help others meet their mission requirements;
- Data analyses and resulting decisions are aided by a broader information resource base when data from others can be aggregated without significant manipulation;
- Standards will be augmented as participation grows and needs expand;
- Time required to specify and complete data exchange agreements is reduced; and
- Data management and transactional costs are reduced.

1. Scope

This EDSC standard is a standard that describes data elements and data blocks that are used to exchange attached binary object data and information.

This EDSC standard is applicable to:

- cataloguing and exchanging information about attached binary objects

This EDSC standard defines:

- data elements that describe attached binary objects

2. Normative References

This standard relies on one other standards to make it complete and provide the necessary support. As such users should consider the Normative Standard (references) noted below, integral to the Attached Binary Object Data Standard Standard. These include:

- ISO 15836:2003 Information and documentation - Dublin Core Metadata Element Set

3. Terms and Definitions

For the purposes of this document, the following terms and definitions apply.

3.1

data element

a single unit of data that cannot be divided and still has useful meaning.

[source: Core Reference Model draft (version 1.0)]

Example Data Elements are individual components of an Address, such as City Name, and Zip Code.

3.2

data block

a grouping of related Data Elements that can be used and reused among different information flows.

[source: Core Reference Model draft (version 1.0)]

An example Data Block is Address Identification, which includes the component Data Elements such as City Name, State Name, and Zip Code.

3.3

attached binary object

Reference documents, images, photos, GIS data layers, laboratory materials and other objects attached within the data exchange.

4. Attached Binary Object Data Blocks and Data Elements

This standard specifies a data block that may be used to identify the characteristics and/or catalog an attached digital data item/object. Table 1 specifies the data elements.

Table 1 – Attached Binary Object Data Blocks and Data Elements

Attached Binary Object Data Standard **Attached Binary Object – Data Elements** *Definition*: An object, its exchange characteristics, and cataloging descriptors.			
Data Element Name	**Data Element Definition**	**Notes**	**Format [tbd]**
1 Binary Object Exchange Characteristics *Definition:* The object and its exchange characteristics. These objects are sometimes referred to as BLOBs (Binary Large Objects).			
1.1 Binary Object	The actual binary object.	Examples: Reference document Image Photo GIS data layer Laboratory material Audio recording	
1.2 Binary Object File Name	The text describing the descriptive name used to represent the file, including file extension.	Example: "Chain of Custody.doc"	
1.3 Binary Object Document Type	The general type of the document object as specified for Network Node conformance.	Permitted values: XML an XML document flat a flat text file bin a binary object ZIP a compressed file in ZIP format other an unspecified file format	
1.4 Binary Object Size	Indicates the binary object size in bytes.	May be useful for sender to determine how best to handle file.	

Attached Binary Object Data Standard **Attached Binary Object – Data Elements** *Definition*: An object, its exchange characteristics, and cataloging descriptors.			
Data Element Name	**Data Element Definition**	**Notes**	**Format [tbd]**
2 Binary Object Dublin Core Descriptors *Definition:* The descriptors used to identify and catalog an object. ***Note: The items below are taken from ISO 15836:2003 Information and documentation - Dublin Core Metadata Element Set*** ***Note: There are internationally accepted XML Schemas for the Dublin Core Metadata Elements.***			
2.1 Binary Object Title	A name given to the resource.	Example: “ABC Company NPDES Compliance Sampling Chain of Custody: 12/31/2003” [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Title".]	
2.2 Binary Object Creator	An entity primarily responsible for making the content of the resource.	Examples of Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Creator".]	
2.3 Binary Object Subject	A topic of the content of the resource.	Typically, Subject will be expressed as keywords, key phrases, or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Subject".]	

Attached Binary Object Data Standard **Attached Binary Object – Data Elements** *Definition*: An object, its exchange characteristics, and cataloging descriptors.			
Data Element Name	**Data Element Definition**	**Notes**	**Format [tbd]**
2.4 Binary Object Description	An account of the content of the resource.	Examples of Description include, but are not limited to, an abstract, table of contents, reference to a graphical representation of content, or free-text account of the content. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Description".]	
2.5 Binary Object Publisher	An entity responsible for making the resource available.	Examples of Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Publisher".]	
2.6 Binary Object Contributor	An entity responsible for making contributions to the content of the resource.	Examples of Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Contributor".]	
2.7 Binary Object Date	A date of an event in the lifecycle of the resource.	Typically, Date will be associated with the creation or availability of the resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and includes (among others) dates of the form YYYYMMDD. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Date".]	Date: yyyymm-dd

Attached Binary Object Data Standard **Attached Binary Object – Data Elements** *Definition*: An object, its exchange characteristics, and cataloging descriptors.			
Data Element Name	**Data Element Definition**	**Notes**	**Format [tbd]**
2.8 Binary Object Type	The nature or genre of the content of the resource.	Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCT]). To describe the physical or digital manifestation of the resource, use the Format element. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Type".]	
2.9 Binary Object Content Format	The physical or digital manifestation of the resource.	Typically, Format will include the media-type or dimensions of the resource. Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats). Examples: “image/jpeg”, “text/html”, “video/mpeg”. Potential Example List take from W3C registered MIME Content Types. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Format".]	

Attached Binary Object Data Standard **Attached Binary Object – Data Elements** *Definition*: An object, its exchange characteristics, and cataloging descriptors.			
Data Element Name	**Data Element Definition**	**Notes**	**Format [tbd]**
2.10 Binary Object Identifier	An unambiguous reference to the resource within a given context.	Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Formal identification systems include but are not limited to the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI), and the International Standard Book Number (ISBN). [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Identifier".]	
2.11 Binary Object Source	A reference to a resource from which the present resource is derived.	The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Source".]	
2.12 Binary Object Language	A language of the intellectual content of the resource.	Recommended best practice is to use RFC 3066 [RFC3066], which, in conjunction with ISO 639 [ISO639], defines two- and three-letter primary language tags with optional subtags. Examples include “en” or “eng” for English, “akk" for Akkadian, and “en-GB” for English used in the United Kingdom. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Language".]	

Attached Binary Object Data Standard

Attached Binary Object – Data Elements

Definition: An object, its exchange characteristics, and cataloging descriptors.

Data Element Name	Data Element Definition	Notes	Format [tbd]
2.13 Binary Object Relation	A reference to a related resource.	Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Relation".]	
2.14 Binary Object Coverage	The extent or scope of the content of the resource.	Typically, Coverage will include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range), or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and to use, where appropriate, named places or time periods in preference to numeric identifiers such as sets of coordinates or date ranges. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Coverage".]	
2.15 Binary Object Rights	Information about rights held in and over the resource.	Typically, Rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions may be made about any rights held in or over the resource. [Note This is from ISO 15836:2003 Information and documentation - The Dublin Core metadata element set Element name is "Rights".]	